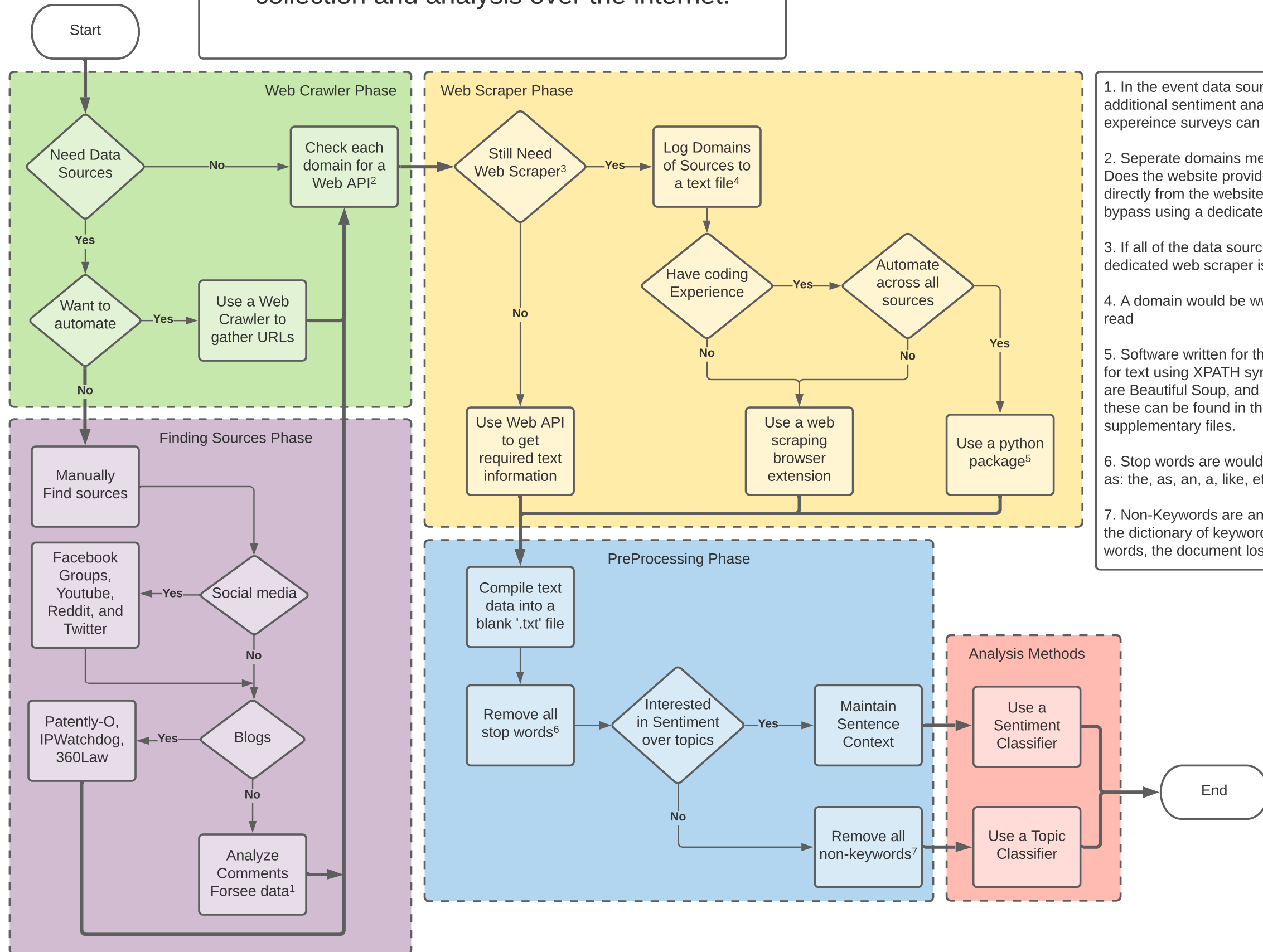


How to conduct customer perception data collection and analysis over the internet.



1. In the event data sources are not found online, additional sentiment analysis of the customer experience surveys can lead to actionable information.

2. Seperate domains mean www.example.com, etc. Does the website provide an interface to request data directly from the website? If so, this can be used to bypass using a dedicated web scraper on the site.

3. If all of the data sources provide an API then a dedicated web scraper is unnecessary.

4. A domain would be www.example.com, this list will be read

5. Software written for the purpose of parsing websites for text using XPATH syntax. Recommended packages are Beautiful Soup, and Scrapy. Further information on these can be found in the literature review section of the supplementary files.

6. Stop words are woulds that have no meaning such as: the, as, an, a, like, etc.

7. Non-Keywords are any word that is not contained in the dictionary of keywords. Without all the meaningful words, the document loses meaning.